


```
source
1. .413
/organism="Homo sapiens"
/db_xref="GDB:3926900"
/db_xref="taxon:9606"
/clone="IMAGE:548384"
/clone_lib="Stratagene neuroepithelium NT2RAMI 937234"
/dev_stage="Ntera-2/RAMi neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: phuescript SK; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. NT2 (Ntera-2/cl.D1) precursor cells induced with Retinoic Acid for 1 week, followed by 3 weeks in mitotic inhibitors (replate 12). Average insert size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' 4 others
```

BASE COUNT 105 a 108 c 71 g 125 t

ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 413; Best Local Similarity 100.0%; Pred. No. 1e-67; Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctaccacatctactcagcagccaccctttgaagttcgtgtgccaatcagtgatcaa 60
|
Db 281 CTACCACATCTACTCCAGCGACCCACCTTTGAAGTTCGTGTGCCANTCAGTGGATCAA 222
|
Qy 61 gtttaagtcagtcagtttaagggaggagaagaggttatccttcaggggggtaccag 120
|
Db 221 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGAGGTATACCTTTCAGGGGGCTACCAG 162
|
Qy 121 acagttctcaactgttgaaggaggagaagaaacccagtcgaatgaattcaatgaatt 180
|
Db 161 ACAGTGTCTCAACTGTGTTAAGGAGGAGAGAAACCCAGTCANTGAATTCAGTGAATTT 102
|
Qy 181 ctggaaactccattgaagttgtagattgacaggttagtaattgcagtcagtttgatca 240
|
Db 101 CTGGAACTCCATTAAAGTGTAGATTGACAGGTAGTAAATTCATGTCAGTTTGATCA 42
|
Qy 241 ttatgcatcaataagataattatgagtcgtt 273
|
Db 41 TTATGTCATTAAGAGATGAATTTATGAGTGCTT 9
|

RESULT 2
AI306649/c

LOCUS

DEFINITION

q38602.x1 NCI_CGAP_kid5 Homo sapiens cDNA clone IMAGE:1900299 3' similar to gb:L13318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.

---SSION
VERSION
KEYWORDS
SOURCE

AI306649.1 GI:3989720
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 424)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/www-bio.llnl.gov/dbbrp/image/image.html at:
www-bio.llnl.gov/dbbrp/image/image.html
Insert Length: 588 Std Error: 0.00

Seq primer: -40UP from Gibco
High quality sequence stop: 392.
Location/Qualifiers
source
1. .424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1900299"
/clone_lib="NCI_CGAP_kid5"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACTGAGAGAAATTCGGCGCGCAATATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 107 a 115 c 72 g 130 t

ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 424; Best Local Similarity 100.0%; Pred. No. 1e-67; Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctaccacatctactcagcagccaccctttgaagttcgtgtgccaatcagtgatcaa 60
|
Db 283 CTACCACATCTACTCCAGCGACCCACCTTTGAAGTTCGTGTGCCANTCAGTGGATCAA 224
|
Qy 61 gtttaagtcagtcagtttaagggaggagaagaggttatccttcaggggggtaccag 120
|
Db 223 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGAGGTATACCTTTCAGGGGGCTACCAG 164
|
Qy 121 acagttctcaactgttgaaggaggagaagaaacccagtcgaatgaattcaatgaatt 180
|
Db 163 ACAGTGTCTCAACTGTGTTAAGGAGGAGAGAAACCCAGTCAGTGAATTCAGTGAATTT 104
|
Qy 181 ctggaaactccattgaagttgtagattgacaggttagtaattgcagtcagtttgatca 240
|
Db 103 CTGGAACTCCATTAAAGTGTAGATTGACAGGTAGTAAATTCATGTCAGTTTGATCA 44
|
Qy 241 ttatgcatcaataagataattatgagtcgtt 273
|
Db 43 TTATGTCATTAAGAGATGAATTTATGAGTGCTT 11
|

RESULT 3
AA149954/c

LOCUS

DEFINITION

z003h02.s1 Stratagene colon (#937204) Homo sapiens cDNA clone IMAGE:566643 3' similar to gb:L13318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

AA149954
AA149954.1 GI:1721168
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 454)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissole, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, K., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478

JOURNAL
MEDLINE
COMMENT

Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. from Amersham
High quality sequence stop: 376.

FEATURES

Location/Qualifiers
1. .454
/organism="Homo sapiens"
/db_xref="GDB:4591096"
/db_xref="taxon:9606"
/clone="IMAGE:566643"
/clone_lib="Stratagene colon (#937204)"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/note="Organ: colon; Vector: phagescript SK-; Site: 1;
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo 48. T-84 colonic epithelial cell line. Average
Insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
CTCAGATTTTTTTTTTTT 3'."

BASE COUNT 114 a 121 c 76 g 142 t 1 others

Query Match 100.0%; Score 273; DB 9; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaccacatctactccagcaccacacctttgaagtcgtgtgcccacatcagtgatcaa 60
Db 276 CTACCACATCTACTCCAGCAGCCACCTTTGAAGTTCGTGTCCTCAATCAGTGCATCA 217
QY 61 gtttaagtcagtcagtttaagggaggagagaggtttacattcaggggggtaccag 120
Db 216 GTTTAAGTCAGTCAGTTAAGGGGAGGAGGAGGTTTATACCTTCAGGGGGGTACCAG 157
QY 121 acagtggttctcaactggttaagggaggagagaggtttacattcaggggggtaccag 180
Db 156 ACAGTGTTCCTCACTGTTAAGGAGGAGGAGGAGGTTTATACCTTCAGGGGGGTACCAG 197
QY 181 ctggaaactccatttaagtgtagattgagcaggttagtaattgcatgattgtaca 240
Db 96 CTGGAACTTCCTCACTGTTAAGGAGGAGGAGGAGGTTTATACCTTCAGGGGGGTACCAG 37
241 ttagtgcattaaagatgaattattgagtcctt 273
Db 36 TTAGTGCATTAAGAGATGAATTATTGAGTGCTT 4

RESULT 4
AA639493/c 456 bp mRNA linear EST 23-OCT-1997
LOCUS
DEFINITION
ng90c11.s1 NCI_CGAP_C09 Homo sapiens CDNA clone IMAGE:1159604 3'
similar to gb:U13318 PYROVATE DEHYDROGENASE E1 COMPONENT, ALPHA
SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.

ACCESSION
AA639493
VERSION
AA639493.1 GI:2563272
KEYWORDS
EST.
SOURCE
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 456)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Ilian Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 424.

FEATURES

Location/Qualifiers
1. .456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1159604"
/clone_lib="NCI_CGAP_C09"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo (Soares4
)."

BASE COUNT 114 a 121 c 76 g 144 t 1 others

Query Match 100.0%; Score 273; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaccacatctactccagcaccacacctttgaagtcgtgtgcccacatcagtgatcaa 60
Db 279 CTACCACATCTACTCCAGCAGCCACCTTTGAAGTTCGTGTCCTCAATCAGTGCATCA 220
QY 61 gtttaagtcagtcagtttaagggaggagagaggtttacattcaggggggtaccag 120
Db 219 GTTTAAGTCAGTCAGTTAAGGGGAGGAGGAGGTTTATACCTTCAGGGGGGTACCAG 160
QY 121 acagtggttctcaactggttaagggaggagagaggtttacattcaggggggtaccag 180
Db 159 ACAGTGTTCCTCACTGTTAAGGAGGAGGAGGAGGTTTATACCTTCAGGGGGGTACCAG 240
QY 181 ctggaaactccatttaagtgtagattgagcaggttagtaattgcatgattgtaca 240
Db 99 CTGGAACTTCCTCACTGTTAAGGAGGAGGAGGAGGTTTATACCTTCAGGGGGGTACCAG 40
QY 241 ttagtgcattaaagatgaattattgagtcctt 273
Db 39 TTAGTGCATTAAGAGATGAATTATTGAGTGCTT 7

RESULT 5
AI799955/c
LOCUS

DEFINITION
AI799955
wc41d03.x1 NCI_CGAP_Pr28 Homo sapiens CDNA clone IMAGE:2321189 3'
similar to gb:U13318 PYROVATE DEHYDROGENASE E1 COMPONENT, ALPHA
SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.

ACCESSION
AI799955.1 GI:5365427
VERSION
AI799955.1

KEYWORDS
EST.

SOURCE
human.

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 456)

AUTHORS
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: ccapbs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmerit-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 606 Std Error: 0.00
Seq primer: -400P from Gibco.
Location/Qualifiers

FEATURES
source

1. .456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2321189"
/clone_lib="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 112 a 120 c 75 g 149 t
ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaccacatctctcagcagccacccttttgaaattcgtggtgccaatcagtgatcaa 60
DB 287 CTACCACTACTCTCAGCGACCCACCTTTTGAAGTTCGTGGTCCCAATCAGTGATCAA 228
QY 61 gtttaagtcagtcagttaaaggaggagagaggttatcacctcagggggtaccag 120
DB 227 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGGTTATACCTTCAGGGGGTACCAG 168
QY 121 acagtgcttccaaactgttaaggaggagagaaacccagtcgaatgaaatt 180
167 ACAGTGTTCTCACTTGTTGTTAAGGAGGAGAGAAACCCAGTCAATGAAATTCATGAAAT 108
QY 181 ctggaaacttccattaaagtgtagatgagcaggttagtaattgcatgagttgtaca 240
DB 107 CTTGGAACCTTCATTAAAGTGTGTAGATTGAGCAGGTAGTAATTGCATGCTTTGTACA 48
QY 241 ttatgcatcattaaagataattattgagtgctt 273
DB 47 TTATGTCATTAAAGATGAATATTAGTGCTT 15

RESULT 6
BE222568/c
LOCUS
DEFINITION
hu48b05.x1 NCI_CGAP_Brn41 Homo sapiens cDNA clone IMAGE:3173265 3', similar to gb:LI3318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT, SOMATIC (HUMAN));, mRNA sequence.
ACCESSION
BE222568
VERSION
BE222568.1 GI:8909886
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 463)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BrGAP), Tumor Gene Index
Unpublished (1998)

JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Email: ccapbs-remail.nih.gov
Tissue Procurement: Robert Jenkins, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D. and M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -400P from Gibco.
Location/Qualifiers

FEATURES
source

1. .463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3173265"
/clone_lib="NCI_CGAP_Brn41"
/tissue_type="oligodendrogloma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACCAATCTGAAGTGGAGCGGCATCAGTCTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 114 a 123 c 76 g 150 t
ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaccacatctctcagcagccacccttttgaaattcgtggtgccaatcagtgatcaa 60
DB 282 CTACCACTACTCTCAGCGACCCACCTTTTGAAGTTCGTGGTCCCAATCAGTGATCAA 223
QY 61 gtttaagtcagtcagttaaaggaggagagaggttatcacctcagggggtaccag 120
DB 222 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGGTTATACCTTCAGGGGGCTACCAG 163
QY 121 acagtgcttccaaactgttagatgagcaggttagtaattgcatgagttgtaca 180
DB 162 ACAGTGTTCTCACTTGTTGTTAAGGAGGAGAGAAACCCAGTCAATGAAATTCATGAAAT 103
QY 181 ctggaaacttccattaaagtgtagatgagcaggttagtaattgcatgagttgtaca 240
DB 102 CTTGGAACCTTCATTAAAGTGTGTAGATTGAGCAGGTAGTAATTGCATGCTTTGTACA 43
QY 241 ttatgcatcattaaagataattattgagtgctt 273
DB 42 TTATGTCATTAAAGATGAATATTAGTGCTT 10

RESULT 7
AA993771/c
LOCUS
DEFINITION

AA993771
ct65h04.s1 Soares-testis_NHT Homo sapiens cDNA clone IMAGE:1621687 3', similar to gb:LI3318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT, SOMATIC (HUMAN));, mRNA sequence.
ACCESSION
AA993771
VERSION
AA993771.1 GI:3180316
KEYWORDS
EST.
SOURCE
human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 478)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 Ph.D.
 ' cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.lnl.gov/nbrp/image/image.html
 Insert Length: 771 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 362.
 Location/Qualifiers
 1. 478
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone IMAGE:1621687"
 /clone lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pYT3D-Pac (Pharmacia) with a modified
 polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACCAATCTGTAAGTGGAGCGGCCGCCCAATTTTTTTTTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pYT3 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 119 a 126 c 81 g 152 t
 BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 273; DB 9; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.le-67;
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ctaccacatctactcagcgaccacacctttgaagttctgtgtgcgaatcagtgatcaa 60
 283 CTACCAACATCTACTCAGCGACCCACCTTTTGAAGTTCGTGTGGTGCCTATCCTTTCAGGGGGCTACCAAG 164
 QY 61 gtttaagtcagtcagtttaaggggaggaagagagattatccttcaggggggtaccag 120
 223 GTTTAAGTCAGTCAGTTAAGGGAGGAGCAGAGAGGCTTATACCTTTCAGGGGGCTACCAAG 164
 QY 121 acagtgcttccaacttgggttaaggggggaagaaaccacgaatgaattccaatgaatt 180
 163 ACAGTGTCTCAACTTGGTTAAGGAGGAGAAACCCAGTCATGAAATTCATGAAT 104
 QY 181 cttggaacttcattaaagtgtgtagattgagcaggtagtaattgcagtcagttgtaca 240
 103 CTTGGAACCTTCCTTAAGTGTGTAGATTGACGAGGTAGTAATTGCATGTCAGTTGTGACA 44
 QY 241 ttatgtcattaaagatgaattatgagtcctt 273
 Db 43 TTAGTGCATTAAAGATGAATTAATTGAGTGCTT 11
 RESULT 8
 AI493781/c AI493781 linear EST 18-MAR-1999
 LOCUS 484 bp mRNA
 DEFINITION q293d05.x1 Soares pregnant.uterus_NBRPU Homo sapiens cDNA clone
 3; similar to gb:U13318 PRYRUATE DEHYDROGENASE E1

VERSION	AA614036.1	GI:2466170
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 485)	
TITLE	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapsb-femail.nih.gov Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D. cDNA Library Arraying: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html Insert length: 1561 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 322. Location/Qualifiers 1. .485 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="NCI-CCAP_AAL" /tissue_type="adrenal adenoma" /lab_host="SOLR (kanamycin resistant)" /note="Organ: adrenal gland; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dr. Two pooled bulk adrenal adenomas. 5' adaptor sequence: 5' GAATTCGCACG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.6 kb." 123 a 129 c 83 g 150 t	
FEATURES	source	
BASE COUNT		
ORIGIN		

Query Match	100.0%	Score 273;	DB 9;	Length 485;
Best Local Similarity	100.0%;	Pred. No. 1.1e-67;		
Matches 273;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ctaccacatctactccagcgaccacacttttgaagttcgtggtggccaaatcagtggatcaa	60	
	280	CTACCACATCTACTCCAGGGACCACCTTTTGAAGTTCTGTTGGTCCCAATCAGTGGATCAA	221	
QY	61	gtttaagtcagtcagttaaagggggagagagaggttatcaccttcaggggggtaccacag	120	
Db	220	GTTTAAAGTCAGTCAGTTTAAAGGGGAGGAGAGAGGTTTATACCTTTCAGGGGCTACCAG	161	
QY	121	acagtgttctcaacttggcttaagggagaaagaaacccagtcagatgaattcgaataaatt	180	
Db	160	ACAGTGTCTCAACTTGGTTAAGGAGGAAGAAACCCAGTCATGAAATTCATGAAATTT	101	
QY	181	cttgaaacttcacttaagttgttagattgagcaggtagtaattgcatgcagtttgtaca	240	
Db	100	CTTGGNAACCTCCATTAAGTGTGTAGATTGAGCAGGTAGTAATTTCATGCAGTTTGTACA	41	
QY	241	ttagtgcaattaaaagatgaattattgagtgctt	273	
Db	40	TTAGTGCATTTAAAGATGAATTAATTGAGTGCTTT	8	

RESULT	10
AI283292/c	
LOCUS	AI283292
DEFINITION	495 bp mRNA linear EST 23-NOV-1998 ql3f08.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:1964679 3' similar to gb:L13318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA

ACCESSION	SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.
VERSION	A1283292
KEYWORDS	A1283292.1 GI:3921525
SOURCE	EST.
ORGANISM	Human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 495) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Seq primer: -40UP from Gibco High quality sequence stop: 374. Location/Qualifiers 1. .495 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="IMAGE:1964679" /clone_lib="NCI-CGAP_Ov23" /tissue_type="tumor, 5 pooled (see description)" /lab_host="DHL0B" /note="Organ: ovary; Vector: PCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"
BASE COUNT	121 a 136 c 90 g 147 t 1 others
ORIGIN	

	Query Match	100.0%;	Score 273;	DB 9;	Length 495;
	Best Local Similarity	100.0%;	Pred. No. 1.le-67;		
	Matches 273;	Conservative 0;	Indels 0;	Gaps 0;	
Qy	1	ctaccacatctactccagcaccacacatttgaagtctggtggccaatcagtgatcaaa	60		
Db	273	CTACCACATCTACTCCAGCAGCACCCCTTTTGAAGTTCGTGGTGCCAATCAGTGGATCAA	214		
Qy	61	gtttaagtcagtcagttaaaggggagggagagaggtattaccttcaggggggtaccag	120		
Db	213	GTTTAAAGTCAGTCAGTTAAAGGGAGGAGAGAGGTTATACCTTCAGGGGGCTACCAG	154		
Qy	121	acagtggtctccaactgttcaaggaggagaagaaaccagtcgaatgaaattcaatgaatt	180		
Db	153	ACAGTGTCTCAACTTGGTTAAGGAGGAGCAAAAACCCAGTCATGAATTCATATGAAATT	94		
Qy	181	cttggaaacttcattaaagtgtagattgagcaggtagtagtaattgcagctttgtaca	240		
Db	93	CTTGGAAACNCTCCATTAAGTGTGTAGATTGAGCAGGTAGTAATTGTCATGTCAGTTGTGACA	34		
Qy	241	ttagtgcatataaagatgaattattgagtgctt	273		
Db	33	TTAGTGCATTAAAGAATGAATTTATTCAGTGCCTT	1		

RESULT 11	AA687297/c	AA687297	497 bp	mRNA	linear	EST 24-DEC-1997
LOCUS		nv61b08.s1				
DEFINITION		NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1234263 3' similar to gb:U13318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA				

Mon Aug 19 09:13:27 2002

similar to qb:L1318 PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA
SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.

ACCESSION AA687297
VERSION AA687297.1 GI:2675488
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 497)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov

unknown library type
Insert Length: 800 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 441.

Location/Qualifiers

1. 497
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1234263"
/clone_lib="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-GTGTACCAATCTGAAGTGGAGGCGCCCTCATTTTTTTTTTTT-3'
) Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 118 a 129 c 81 g 169 t
ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/ 1 ctaccacatctactccagcaccaccccttttgaagttcgtggtgccaatcagtgatcaa 60
Db 295 CTACCACATCTACTCCAGGAGCCACCTTTTGAAGTTCGTGTGCGCAATCAGTGGATCAA 236
QY 61 gtttaagtcagtcagtttaagggaggagagagaggttatccttcagggggctaccag 120
Db 235 GTTTAAGTCAGTCAGTTAAGGGGAGGAGAGGAGGTTATACCTTCAGGGGGCTACCAG 176
QY 121 acagtgcttcaactgtgttaaggaggagagaaacccagtcacatgaattcaatgaatt 180
Db 175 ACAGTGTTCTCACTTGTGTTAAGGAGGAGAAACCCAGTCAATGAATTCATGAATTT 116
QY 181 ctgggaacttccttaagtggtgattgagcaggttagtaattgcatcaggtttgtaca 240
Db 115 CTTGGAACCTCCATTAAGTGTGTAGATTGAGCAGGTAGTATTGATGTCAGTTGTACA 56
QY 241 ttagtgcattaaaagaatgaattatgagtgctt 273
Db 55 TTAGTGCATTAAAGATGAATTTATTGAGTGCTT 23

RESULT 12
AA847355 503 bp mRNA linear EST 31-WAR-1998
LOCUS 0682c07.s1 NCI_CGAP_Col12 Homo sapiens cDNA clone IMAGE:1418124

ACCESSION AA847355
VERSION AA847355.1 GI:2933873
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 503)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov

Emmert-Buck, M.D., Ph.D.
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
cDNA Library Preparation: Stratagene, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1140 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 290.

Location/Qualifiers

1. 503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1418124"
/clone_lib="NCI_CGAP_Col12"
/sex="mixed"
/tissue_type="colon tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: colon; Vector: Bluescript SK-; Site.1: EcoRI
; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Pooled colon tumors. 5' adaptor sequence: 5'
GAATTCGGCAGGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.2 kb."

BASE COUNT 127 a 132 c 95 g 149 t
ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaccacatctactccagcaccaccccttttgaagttcgtggtgccaatcagtgatcaa 60
Db 281 CTACCACATCTACTCCAGGAGCCACCTTTTGAAGTTCGTGTGCGCAATCAGTGGATCAA 222
QY 61 gtttaagtcagtcagtttaagggaggagagaggttatccttcagggggctaccag 120
Db 221 GTTTAAGTCAGTCAGTTAAGGGGAGGAGAGGAGGTTATACCTTCAGGGGGCTACCAG 162
QY 121 acagtgcttcaactgtgttaaggaggagagaaacccagtcacatgaattcaatgaatt 180
Db 161 ACAGTGTTCTCACTTGTGTTAAGGAGGAGAAACCCAGTCAATGAATTCATGAATTT 102
QY 181 ctgggaacttccttaagtggtgattgagcaggttagtaattgcatcaggtttgtaca 240
Db 101 CTTGGAACCTCCATTAAGTGTGTAGATTGAGCAGGTAGTATTGATGTCAGTTGTACA 42
QY 241 ttagtgcattaaaagaatgaattatgagtgctt 273
Db 41 TTAGTGCATTAAAGATGAATTTATTGAGTGCTT 9

RESULT 13
AA847355/c AI313205/LOCUS 512 bp mRNA linear EST 03-FEB-1999

LOCUS	BE672988	519 bp	mRNA	linear	EST 08-SEP-2000							
DEFINITION	7026c06.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3248362 3' similar to gb:U3318 PYROVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT, SOMATIC (HUMAN);, mRNA sequence.											
ACCESSION	BE672988											
VERSION	EST.											
KEYWORDS	BE672988.1	GI:10033529										
SOURCE	human.											
ORGANISM	Homo sapiens											
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
AUTHORS	1 (bases 1 to 519)											
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.											
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index											
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: -40UP from Gibco High quality sequence stop: 460.											
FEATURES	Location/Qualifiers											
source	1..519											
	/organism="Homo sapiens"											
	/db_xref="taxon:9606"											
	/clone="IMAGE:3248362"											
	/clone_lib="NCI_CGAP_Pr28"											
	/sex="male"											
	/dev_stage="adult"											
	/lab_host="DH10B"											
	/note="Organ: prostate; Vector: pT773D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."											
BASE COUNT	128 a	137 c	91 g	163 t								
ORIGIN												
Query Match	100.0%; Score 273; DB 10; Length 519;											
Best Local Similarity	100.0%; Pred. No. 1.le-67;											
Matches 273; Conservative.	0; Mismatches 0; Indels 0; Gaps 0;											
Qy	1	ctaccatctactccagcaccaccccttttgagttcgtggtgcceaatcagtgatcaa	60									
Db	282	CTACCACATCTACFCCAGCACCACCTTTTGAAGTTCGTGGTGCCCAATCAGTGGATCAA	223									
Qy	61	gttaagtcagtcagttaaaggaggagagagattataccttcaggggggtaccag	120									
Db	222	GTTTTAAGTCAGTCAGTTTAAAGGGAGGAGAGGAGAGGTTATACCTTCAGGGGGCTACCAG	163									
Qy	121	acagtggttccaacttggttaaggaggagaagaaaccagtcgaatgaaattcaatgaatt	180									
Db	162	ACAGTGTCTCAACTTGGTTAAGGAGGAGAGAAACCCAGTCATGAATGAATGAAT	103									
Qy	181	cttggaactccattaaagtgtgtagattgagcaggtagtaattgcagcttggtaca	240									
Db	102	CTTGAACATCTCCATTAAAGTCTGTAGATTGAGCAGGTAGTAATTGCATGCAGTTGTACA	43									
Qy	241	ttagtgcattaaaagatgaattatttgagtgc	273									
Db	42	TTAGTGCATTAAAGATGAATTAATTGATGCTTT	10									

Search completed: August 16, 2002, 13:08:49
Job time: 5174 sec

```

FEATURES
source
Location/Qualifiers
1. .537
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2151813"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: p773D-pac (Pharmacia) with a
modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
tcttaccagctcgaatgctggagcgccgcctctttttttttttttttttttttt
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' AATTCACGACTAAT 3' and 5' ATTACGATG 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified p773 vector. Library constructed by Bob
Barstead."

```

BASE COUNT	133 a	147 c	95 g	162 t	Barstead.*
ORIGIN					
Query Match	100.0%	Score 273;	DB 9;	Length 537;	
Best Local Similarity	100.0%;	Pred. NO. 1.le-67;			
Matches 273;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ctaccacatctactccagcagccacaccttttgaagttcgtgggtgccaatcagtggaatcaa	60		
DB	280	CTACCCACATCTACTCCAGCGACCCACCTTTTGAAGTTTCGTGGTGCCTCAATCAGTGGATCAA	221		
QY	61	gtttaagtcagtcagttaaagggggggagaagagaggtttatacttcaggggggtaccag	120		
DB	220	GTTTAAGTCAGTCAGTCTAAGGGGAGGAGGAAGGAGGTTTATCCTTCAGGGGGCTACCCAG	161		
QY	121	acaagttctccaacttggttaaggagggaagaaaccccaagtcgaatgaatcaatgaatt	180		
DB	160	ACAGTGTCTCAACTTGGTTAAGGAGGAAGAAACCCAGTCAATGAATTCATGGAATTT	101		
QY	181	cttggaaacttcataatgaagttagattgagcaggtagtaattgcattgcagtttgtaca	240		
DB	100	CTTGGAAACTTCCATTAAAGTGTGTAGATTTGACGAGGTAGTAATTTCATGCGAGTTGTGACA	41		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 11:49:05 ; Search time 46.68 Seconds
(without alignments)
1436.546 Million cell updates/sec

Title: US-09-761-579-2_COPY_1200_1472

Perfect score: 273
Sequence: 1 ctaccacatctactccagcg.....agatgaattattgagtcgtt 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

tal number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	32.2	11.8	7218	1	US-08-232-463-14
2	31.6	11.6	394	2	US-08-623-906A-7
3	31.6	11.6	2938	2	US-08-343-443B-3
4	31.4	11.5	1004	2	US-08-540-804-7
5	31.4	11.5	1004	2	US-08-218-265-7
6	31.4	11.5	1004	3	US-08-521-872-7
7	31.4	11.5	1004	4	US-08-590-399-7
8	30	11.0	1029	3	US-08-978-456-1
9	30	11.0	1029	4	US-09-369-700-1
10	30	11.0	1925	2	US-08-553-436A-1
11	30	11.0	3336	2	US-08-977-554-7
12	30	11.0	3336	3	US-08-978-456-7
13	30	11.0	3336	4	US-09-225-967-7
14	30	11.0	3336	4	US-09-369-700-7
15	30	11.0	3336	4	US-09-227-806-7
16	30	11.0	3336	4	US-09-491-916-5
17	29.4	10.8	1147	3	US-08-755-587-42
18	29.4	10.8	1147	3	US-08-676-967-2
19	29.4	10.8	2277	1	US-08-676-974-2
20	29.4	10.8	2277	2	US-09-098-487-2
21	29.4	10.8	3324	2	US-08-820-170A-33
22	29.4	10.8	3324	3	US-09-055-699-33
23	29.4	10.8	3324	4	US-09-273-565-33
24	29.4	10.8	3324	4	US-09-565-538-33
25	29	10.6	3362	3	US-09-344-519-1
26	28.6	10.5	1359	1	US-07-618-312A-1
27	28.6	10.5	1359	1	US-08-110-786A-7

28 28.6 10.5 1359 1 US-08-280-228-1 Sequence 1, Appl
29 28.6 10.5 1858 1 US-08-668-381A-6 Sequence 6, Appl
c 30 28.4 10.4 1117 2 US-08-960-022-1 Sequence 1, Appl
c 31 28.4 10.4 19011 1 US-08-310-356-36 Sequence 36, Appl
c 32 28.4 10.4 19557 5 PCT-US92-06300-1 Sequence 1, Appl
c 33 28.2 10.3 1186 2 US-08-731-722-5 Sequence 5, Appl
c 34 28.2 10.3 1218 2 US-08-731-722-4 Sequence 1, Appl
c 35 28.2 10.3 3926 2 US-08-731-722-1 Sequence 1, Appl
c 36 28.2 10.3 3926 2 US-08-731-722-1 Sequence 1, Appl
c 37 28.2 10.3 3926 2 US-08-731-722-2 Sequence 2, Appl
c 38 28.2 10.3 3926 2 US-08-731-722-2 Sequence 2, Appl
c 39 28.2 10.3 3933 2 US-08-731-722-3 Sequence 3, Appl
c 40 28.2 10.3 3933 2 US-08-731-722-3 Sequence 3, Appl
c 41 27.8 10.2 1501 3 US-08-993-359-21 Sequence 21, Appl
c 42 27.8 10.2 2160 1 US-08-700-749A-8 Sequence 8, Appl
c 43 27.8 10.2 2160 3 US-09-020-684-8 Sequence 8, Appl
c 44 27.8 10.2 2160 3 US-09-020-467-8 Sequence 8, Appl
c 45 27.8 10.2 2160 3 US-09-020-685-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)836-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT29pt-F1s
; US-08-232-463-14

```

NAME/KEY: misc_feature
LOCATION: 350..394
OTHER INFORMATION: /note= "Nucleotides 350-394 are unique flanking sequence"
US-08-623-906A-7

Query Match      11.8%; Score 32.2; DB 1; Length 7218;
Best Local Similarity 4.3%; Pred. No. 0.94; Indels    0; Gaps    0;
Matches 10; Conservative 129; Mismatches          92;

QY   30 tgaagtggtgctcgaatcaagtgcatttaagaacttcacaggcgtagggagaga 89
Db -1450 TAGAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1391
QY   90 agagagggtataccgtcagggggtaccagacagtgcttccaacttggttaagdgaa 149
Db 1390 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1331
QY   150 gaaaaccagtcgaatgaattcaaatcttggaaccttcacattaagtgtgatatt 209
Db 1330 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1271
QY   210 gagcagtagtaattcatcgatgctttgttacattagtcattaaagaatgaa 260
Db 1270 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1220

JULI  2
us-08-623-906A-7
Sequence 7, Application US/08623906A
Patent No. 5874217
GENERAL INFORMATION:
APPLICANT: Stevenson, Tamara
APPLICANT: Dvorak, Jan
APPLICANT: Halverson, Joy
TITLE OF INVENTION: Microsatellite Sequences for Canine Genotyping
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLHRH, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/623,906A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-62282/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..233
OTHER INFORMATION: /note= "Nucleotides 1-233 are unique flanking sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 234..349
OTHER INFORMATION: /note= "Nucleotides 234-349 are repeat sequence"
FEATURE:

```


us-09-761-579-2_copy_1200_1472.rni

Mon Aug 19 09:13:25 2002

```

GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: ribG
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-1

Query Match 11.0%; Score 30; DB 3; Length 1029;
Best Local Similarity 54.5%; Pred. No. 2.2;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 139 ttaaggaggagaagaaacccagtcgaatgaattcgaattcgtggaactccattaa 198
Db 596 TTCAAGATGGAAGAAACCCCTATAAAGTAATATGCTAAGTCTGGGAATATTCATTTTA 655

199 gtgtgtagtgagcaggtagtagtgcagtcgagttgtacattagtcca 248
Db 656 ATCAGCAATTTATCAAGATGAATCAACACCAATTGGATATATACTGAA 705

RESULT 10
US-08-553-436A-1/C
Sequence 1, Application US/08553436A
Patent No. 5866790
GENERAL INFORMATION:
APPLICANT: HESSE, Holger
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE
CONCENTRATION
TITLE OF INVENTION: PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE
CONCENTRATION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,436A
FILING DATE: 17-NOV-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/01671
FILING DATE: 20-MAY-1994
PRIOR APPLICATION NUMBER: DE P 4317596.1
FILING DATE: 24-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weillman, Edward
REGISTRATION NUMBER: 24,735
US-09-369-700-1

Query Match 11.0%; Score 30; DB 3; Length 1029;
Best Local Similarity 54.5%; Pred. No. 2.2;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 139 ttaaggaggagaagaaacccagtcgaatgaattcgaattcgtggaactccattaa 198
Db 596 TTCAAGATGGAAGAAACCCCTATAAAGTAATATGCTAAGTCTGGGAATATTCATTTTA 655

199 gtgtgtagtgagcaggtagtagtgcagtcgagttgtacattagtcca 248
Db 656 ATCAGCAATTTATCAAGATGAATCAACACCAATTGGATATATACTGAA 705

RESULT 9
US-09-369-700-1
Sequence 1, Application US/09369700
Patent No. 6280735
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: ribG
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

```

;
FILING DATE: 19-FEB-1997

1-924-815-80-20

US-08-978-430-1

APPLICANT: Warren, Richard L.
APPLICANT: Traini, Christopher M.
APPLICANT: Wang, Min
APPLICANT: Jaworski, Deborah D.
APPLICANT: Mooney, Jeffrey
APPLICANT: Debouck, Christine
APPLICANT: Zhong, Yiyl
APPLICANT: Black, Michael
TITLE OF INVENTION: ribA
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/227,806
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/977,554
FILING DATE:
APPLICATION NUMBER: PCT/US97/02318
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Q. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-227-806-7

Query Match 11.0%; Score 30; DB 4; Length 3336;
est Local Similarity 54.5%; Pred. No. 3.4;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
Qy 139 ttaaggagggaagaaaccccggtcaatgaaattcattcttggaacttcatttaa 198
Db 596 TTCAGATGGAAGAAACCCCTATAAAGTATATTGCTAAGTCTGGGAATATTCATTTA 655
Qy 199 gtgtgtagattgagcaggtagtgatgcagtcagttgttacattagtgca 248
Db 656 ATCAGCAAAATTATCAAGATGATCAACACCAATTTGGATATATATACTGAA 705

Search completed: August 16, 2002, 13:41:11
Job time: 6726 sec

us-09-761-579-2-copy_1200_1472.rng

Mon Aug 19 09:13:24 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 16, 2002, 12:35:24 ; Search time 211.74 Seconds
(without alignments)
2213.647 Million cell updates/sec

Title: US-09-761-579-2_COPY_1200_1472
Perfect score: 273
Sequence: 1 ctaccacatctactccgcg.....agatgaattattgagtcct 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapert 1.0

Searched: 1736436 seqs, 858457221 residues
tal number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	79.5	229	16 AAT21382	Human gene signatu
2	215.4	78.9	275	22 AAL23550	Human breast cance
3	210.4	77.1	311	22 ABL14688	Human breast cance
4	34.6	12.7	7516	22 ABL32060	Human immune syste
5	33.8	12.4	27179	22 AAS34556	Human DNA for a no
6	33.6	12.3	113515	22 ABL34175	Human immune syste
7	33.2	12.2	2700	23 AAS65647	DNA encoding novel
8	33.2	12.2	3613	23 ABL16660	Drosophila melanog
9	33.2	12.2	4356	23 ABL16478	Drosophila melanog

10	32.6	11.9	1464	21 AAC48192	Arabidopsis thalia
11	32.6	11.9	18890	23 ABL13818	Drosophila melanog
12	31.8	11.6	878	21 AAC69616	Human secreted pro
13	31.8	11.6	6899	24 ABL33894	Human immune syste
14	31.6	11.6	394	18 AAT65070	Canine genomic mic
15	31.6	11.6	394	20 AAX17807	Microsatellite rep
16	31.6	11.6	2938	14 AAQ50644	Human Hum-Fli-1 ge
17	31.6	11.6	2954	22 AAH02915	Human shear stress
18	31.4	11.5	1004	18 AAT59906	yeast transcriptio
19	31.2	11.4	8323	24 ABL32058	Human immune syste
20	31	11.4	3227	22 AAS03283	C. elegans cDNA en
21	31	11.4	3709	22 AAS03282	C. elegans genomic
22	31	11.4	1503900	22 AAK95240	Human neuregulin-1
23	31	11.4	1503900	22 AAK96733	Human neuregulin-1
24	30.8	11.3	3303	21 AAA70147	Plasmodium falcipa
25	30.8	11.3	63563	22 AAF28546	Genomic fragment #
26	30.6	11.2	240	22 AAI91712	Human polynucleoti
27	30.6	11.2	2784	21 AAC50478	Arabidopsis thalia
28	30.6	11.2	3072	21 AAC47196	Arabidopsis thalia
29	30.6	11.2	3594	21 AAC50212	Arabidopsis thalia
30	30.6	11.2	6665	22 AAS45299	Chemically pretrea
31	30.6	11.2	6665	24 ABL32083	Human immune syste
32	30.6	11.2	13158	22 AAS01100	Human immune/haema
33	30.4	11.1	1587	22 AAS01100	Corn sterol acyltr
34	30.4	11.1	5874	20 AAX59765	cDNA encoding huma
35	30.4	11.1	17142	22 AAS45498	Chemically pretrea
36	30.4	11.1	17142	24 ABL34107	Human immune syste
37	30.2	11.1	547	21 AAZ87243	Partial cDNA clone
38	30.2	11.1	695	21 AAZ87236	cDNA clone #2 enco
39	30.2	11.1	1726	22 AAL00281	Human reproductive
40	30.2	11.1	3753	23 AAS68937	DNA encoding novel
41	30.2	11.1	6153	22 AAH18670	Human cDNA sequen
42	30.2	11.1	8031	22 AAL03290	Human reproductive
43	30	11.0	304	22 AAK64372	Human immune/haema
44	30	11.0	1029	20 AAX59909	Pyrimidine deamina
45	30	11.0	1094	21 AAF21717	Human breast and o

ALIGNMENTS

RESULT 1
AAT21382
ID AAT21382 standard; cDNA to mRNA; 229 BP.
XX AAT21382;
XX AAT21382;
AC AAT21382;
DT 01-AUG-1996 (first entry)
XX Human gene signature HUMGS02746.
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX Homo sapiens.
XX WO9514772-Al.
XX 01-JUN-1995.
XX 11-NOV-1994; 94WO-JP01916.
XX 12-NOV-1993; 93JP-0355504.
XX (MATS/) MATSUBARA K.
XX (OKUB/) OKUBO K.
XX Matsubara K, Okubo K;
XX WPI; 1995-206931/27.
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
XX
XX Claim 1: Page 872-873; 2245pp; Japanese.
XX
XX A single-stranded DNA (or its complementary strand or the corresp.
XX double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX given in AAT19001-T26837 and which is able to hybridise to part of
XX human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX sequences were obtained from 3'-directed cDNA libraries prepared
XX from various human tissues; synthesis of cDNA was initiated from the
XX 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
XX is constructed so as to reflect accurately the relative abundance of
XX different mRNAs in the particular tissue from which it was derived.
XX The appearance frequency of a given GS in a cDNA library can be
XX determined (esp. using primers and probes derived from the GS
XX sequences) as a means of diagnosing abnormal cell function or for
XX recognising different cell types.

Query Match 79.5%; Score 217; DB 16; Length 229;
Best Local Similarity 99.1%; Pred. No. 4.6e-59;
Matches 217; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 55 gatcaagtttaagtcagtcagtttaagggaggaggaagggagaggtttacattcagggggc 114
DB 1 gatcaagtttaagtcagtcagtttaagggaggaggaagggagaggtttacattcagggggc 60
QY 115 taccagacagtggtctcaacttggttaagggaggaagaaacccagtcgaatgaattcaat 174
DB 61 taccagacagtggtctcaacttggttaagggaggaagaaacccagtcgaatgaattcaat 120
QY 175 gaattcttggaacttcattcaattgaagtgtagattgagcaggtagtaattgcatgagtt 234
DB 121 gaattcttggaacttcattcaattgaagtgtagattgagcaggtagtaattgcatgagtt 180
QY 235 tgcattatgagtcattaaagatgaattattgagtgctt 273
DB 181 tgcattatgagtcattaaagatgaattattgagtgctt 219

RESULT 2

AAL23550
AAL23550 standard; cDNA; 275 BP.

AAL23550;
XX 07-DEC-2001 (first entry)
XX Human breast cancer expressed polynucleotide 16007.
XX Human; breast cancer; cell marker; cytostatic; ss.
XX Homo sapiens.
XX WO200151628-A2.
XX 19-JUL-2001.
XX 10-JAN-2001; 2001WO-US00798.
XX 14-JAN-2000; 2000US-0176077.
XX 14-MAR-2000; 2000US-0189167.
XX 24-MAR-2000; 2000US-0192099.
XX 29-MAR-2000; 2000US-0193480.
XX 15-MAY-2000; 2000US-0205230.
XX 09-JUN-2000; 2000US-0211315.
XX 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
XX
XX Claim 1: Page 2922; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AAL07544-AAL26789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterising treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cytostatic
XX activity.
XX
XX Sequence 275 BP; 78 A; 59 C; 63 G; 75 T; 0 other;

Query Match 78.9%; Score 215.4; DB 22; Length 275;
Best Local Similarity 93.4%; Pred. No. 1.6e-58;
Matches 225; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ctaccacatctactcagcagccaccctttgaagttctggtgccaatcagtgatcaa 60
DB 29 ctaccacatctactcagcagccaccctttgaagttctggtgccaatcagtgatcaa 88
QY 61 gtttaagtcagtcagtttaagggaggaggaagagaggttacccttcagggggctaccag 120
DB 89 gtttaagtcagtcagttttatggaggaggaatgataggttataccttcagggggctaccat 148
QY 121 acagtggttcactggttaagggagggaagaaacccagtcgaatgaattcaatgaatt 180
DB 149 acagtggttcactggttaagggagggaagaaacccagtcgaatgaattcaatgaatt 208
QY 181 ctggaacattccattcaagtgtagattgagcaggtagtaattgcatgagtttataca 240
DB 209 ctggaacattccattcaagtgtagattgagcaggtagtaattgcatgagtttataca 268
QY 241 t 241
DB 269 t 269

RESULT 3

AAL14688
ID AAL14688 standard; cDNA; 311 BP.

XX
XX AAL14688;
XX
XX 07-DEC-2001 (first entry)
XX Human breast cancer expressed polynucleotide 7145.
XX Human; breast cancer; cell marker; cytostatic; ss.
XX Homo sapiens.
XX WO200151628-A2.
XX 19-JUL-2001.
XX 10-JAN-2001; 2001WO-US00798.
XX 14-JAN-2000; 2000US-0176077.
XX 14-MAR-2000; 2000US-0189167.
XX 24-MAR-2000; 2000US-0192099.
XX 29-MAR-2000; 2000US-0193480.

3d
00-20

PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
PT	
XX	Claim 1; SEQ ID NO 907; 21pp + Sequence Listing; English.
PS	
XX	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences .
CC	
XX	Sequence 4356 BP; 1165 A; 942 C; 1021 G; 1228 T; 0 other;
SQ	

	Query Match	12.2%;	Score 33.2;	DB 23;	Length 4356;
	Best Local Similarity	64.1%;	Pred. No. 2.9;		
	Matches 50;	Conservative	0;	Mismatches 28;	Indels 0;
	Gaps	0;			
y	137	ggttaaggaggagaaacccagtcgaatgaattctctgggaacttcatt	196		
b	1192	ggataaacaataaacacacgaagaattcaagagcaggtttggtaacttgcaat	1251		
y	197	aagtggtgattagattgagca	214		
b	1252	taagctttaattatgaagca	1269		

RESULT 10	
AAC48192	
D	AAC48192 standard; DNA; 1464 BP.
X	
C	AAC48192;
X	18-OCT-2000 (first entry)
X	
X	Arabidopsis thaliana DNA fragment SEQ ID NO: 56594.
E	
X	Hybridisation assay; genetic mapping; gene expression control;
X	protein identification; signal transduction pathway;
X	metabolic pathway; promoter; termination sequence; ss.
X	Arabidopsis thaliana.

[illegible]

PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	08-MAY-1999;	99US-0132488.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135123.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	10-JUN-1999;	99US-0139119.
PR	14-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	02-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	15-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.

PR	22-JUL-1999;	99US-0145085
PR	22-JUL-1999;	99US-0145087
PR	22-JUL-1999;	99US-0145089
PR	22-JUL-1999;	99US-0145192
PR	23-JUL-1999;	99US-0145145
PR	23-JUL-1999;	99US-0145218
PR	23-JUL-1999;	99US-0145224
PR	26-JUL-1999;	99US-0145276
PR	27-JUL-1999;	99US-0145913
PR	27-JUL-1999;	99US-0145918
PR	27-JUL-1999;	99US-0145919
PR	28-JUL-1999;	99US-0145951
PR	28-JUL-1999;	99US-0146386
PR	02-AUG-1999;	99US-0146388
PR	02-AUG-1999;	99US-0146389
PR	03-AUG-1999;	99US-0147038
PR	04-AUG-1999;	99US-0147204
PR	04-AUG-1999;	99US-0147302
PR	05-AUG-1999;	99US-0147192
PR	05-AUG-1999;	99US-0147260
PR	06-AUG-1999;	99US-0147303
PR	06-AUG-1999;	99US-0147416
PR	09-AUG-1999;	99US-0147493
PR	09-AUG-1999;	99US-0147935
PR	10-AUG-1999;	99US-0148171
PR	11-AUG-1999;	99US-0148319
PR	12-AUG-1999;	99US-0148431
PR	13-AUG-1999;	99US-0148365
PR	13-AUG-1999;	99US-0148684
PR	16-AUG-1999;	99US-0149368
PR	17-AUG-1999;	99US-0149175
PR	18-AUG-1999;	99US-0149426
PR	20-AUG-1999;	99US-0149722
PR	20-AUG-1999;	99US-0149723
PR	20-AUG-1999;	99US-0149929
PR	23-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150566
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151066
PR	27-AUG-1999;	99US-0151080
PR	30-AUG-1999;	99US-0151303
PR	31-AUG-1999;	99US-0151438
PR	01-SEP-1999;	99US-0151930
PR	07-SEP-1999;	99US-0152363
PR	10-SEP-1999;	99US-0153070
PR	13-SEP-1999;	99US-0153758
PR	15-SEP-1999;	99US-0154018
PR	20-SEP-1999;	99US-0154039
PR	20-SEP-1999;	99US-0154779
PR	22-SEP-1999;	99US-0155139
PR	23-SEP-1999;	99US-0155486
PR	24-SEP-1999;	99US-0155659
PR	28-SEP-1999;	99US-0156458
PR	29-SEP-1999;	99US-0156596
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157865
PR	08-OCT-1999;	99US-0158029
PR	08-OCT-1999;	99US-0158232
PR	12-OCT-1999;	99US-0158369
PR	13-OCT-1999;	99US-0159293
PR	13-OCT-1999;	99US-0159294
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159330
PR	14-OCT-1999;	99US-0159331
PR	14-OCT-1999;	99US-0159637
PR	14-OCT-1999;	99US-0159638
PR	14-OCT-1999;	99US-0159584
PR	21-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160767

region comprises: (a) a repeated tetranucleotide motif consisting of AAA or TTT and a fourth nucleotide; or (b) a repeated motif selected from AAAG, GAAA, AAAT, TTTC, TTTA, GAAT, GAAG, GAAA, AAAAG, TGC and TTC. The amplification product is size-fractionated to provide a measure of the size of the chromosomal DNA between the primers, where the size of the amplification product is polymorphic for the locus and provides a genotype for the canine. The combined information from multiple loci provides a means of distinguishing individuals, even among inbred dog breeds, for parentage testing, forensic testing and analysis of individual relatedness.

Sequence 394 BP; 96 A; 57 C; 47 G; 194 T; 0 other; XX SQ

Query Match	11.6%	Score 31.6;	DB 20;	Length 394;
Best Local Similarity	55.5%;	Pred. No. 3.6;		
Matches	61;	Conservative	0;	Mismatches 49;
			Indels	0;
			Gaps	0;
QY	147	gaagaaaaaccgcagtcgaatgaattcaatgaattctctggaaacttccattaaagtgtgtag	206	
	73	gtatataaacctccctttctaatgtatgaacaaatttggtatataactctggatgaattg	132	

Qy 207 attgagcaggtagtaattgcagtttgatcattagtgattaaaga 256
|| ||| || | ||| || ||| | ||| ||
nb 133 atcgagaaagaacctattttatatgatttaaagtgaggattatgga 182

Search completed: August 16, 2002, 13:45:32
Job time: 4208 sec

Mon Aug 19 09:13:23 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 16, 2002, 11:46:41 ; Search time 1853.82 Seconds
(without alignments)
3081.715 Million cell updates/sec

Title: US-09-761-579-2_COPY_1200_1472
Perfect score: 273
Sequence: 1 ctaccacatctactccagcg.....agatgattattgagtgtt 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.in.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	273	100.0	367	9	HUMPDHA12
2	273	100.0	1410	9	HUMPYDH
3	273	100.0	1419	9	HUMPDHA
4	273	100.0	1446	9	HUMPDH1B
5	273	100.0	1472	6	AX151747
6	273	100.0	1472	9	HUMPDH1A
7	273	100.0	17072	9	HUMPDHAL
8	273	100.0	195191	2	AC025563
9	270	98.9	1427	9	HSPDHBE1
10	270	98.9	1456	9	HSPDHBE1
11	268.2	98.2	1492	9	HUMPDHA1G
12	267	97.8	1417	9	HSPDHLE1
13	194	71.1	201	11	G13505
14	115	42.1	129	9	S57358
15	90	33.0	149	9	S81593
16	69	25.3	2854	10	RRPDHYE1A
17	67.4	24.7	1446	10	MUSPDHA1A
18	67.4	24.7	2778	10	BC007142
19	67.4	24.7	158565	2	AC106221
20	65.4	24.0	3055	9	HUMPYRDH
21	65.4	24.0	111547	2	AP002332
22	65.4	24.0	146072	9	AC093828
23	63.6	23.3	1344	4	SNPDEHA
24	63	23.1	1170	4	SDPDE1A
25	49	17.9	67	9	S56181
26	47.8	17.5	2349	10	MUSPDHA2A
27	47.8	17.5	58040	2	AC101628
28	47.8	17.5	211301	2	AC100752
29	46	16.8	1616	10	RRPDE1A2
30	46	16.8	2228	10	RNU4125
31	42.2	15.5	154854	2	AC015539
32	42.2	15.5	171589	2	AC073567
33	42.2	15.5	175466	2	AC016419
34	39.8	14.6	163794	2	AC103449
35	39.4	14.4	157511	9	AC074348
36	38.6	14.1	145269	2	AC096563
37	38.4	14.1	168662	9	AC079880
38	38.4	14.1	186343	9	AL137248
39	38.4	14.1	196966	2	AC031975
40	38.2	14.0	39002	9	AF001217
41	38.2	14.0	39494	9	AP001217
42	38.2	14.0	39946	9	AP001220
43	38.2	14.0	109101	2	AC108674
44	38.2	14.0	143710	2	AC036109
45	38.2	14.0	179255	2	AC020636

ALIGNMENTS

RESULT 1

HUMPDHA12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

HUMPDHA12 367 bp DNA linear PRI 08-JAN-1995
Human nuclear-encoded mitochondrial pyruvate dehydrogenase E1 alpha
(pdhE1-a) subunit gene, exon 11.
M27257 J02734 M27166
M27257.1 GI:488487
dehydrogenase; nuclear-encoded mitochondrial EF-G-like protein;
pyruvate dehydrogenase.
12 of 12
Human fetal liver, cDNA to mRNA, clones lambda-PDH[1,5,7 and F];
and leukocyte DNA, clones cpDH[3,4a].
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 9 to 367)
Dahl, H.H., Hunt, S.M., Hutchison, W.M. and Brown, G.K.
The human pyruvate dehydrogenase complex. Isolation of cDNA clones
for the E1 alpha subunit, sequence analysis, and characterization
of the mRNA
J. Biol. Chem. 262 (15), 7398-7403 (1987)
8722349
2 (bases 1 to 14; 168 to 173)

AUTHORS Maragos,C., Hutchison,W.M., Hayasaka,K., Brown,G.K. and Dahl,H.H.
 TITLE Structural organization of the gene for the El alpha subunit of the human pyruvate dehydrogenase complex
 JOURNAL J. Biol. Chem. 264 (21), 12294-12298 (1989)
 MEDLINE 89308653
 COMMENT On May 21, 1994 this sequence version replaced gi:189750.
 Draft entry and computer-readable sequence of [1] kindly provided by H.-H.M.Dahl, 10-APR-1987.

FEATURES
 source
 1. .367
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="leukocyte"
 gene
 join(M29141.1..651..780,M29155.1..1..169,M29156.1..1..76,
 M29157.1..1..190,M29158.1..1..143,M29159.1..1..108,
 M29160.1..1..109,M29161.1..1..172,M29162.1..1..88,
 M29163.1..1..84,M29164.1..1..125,1..367)
 /gene="PDHA1"
 CDS
 join(M29155.1..105..161,M29156.1..9..68,M29157.1..9..182,
 M29158.1..9..135,M29159.1..9..100,M29160.1..9..101,
 M29161.1..9..164,M29162.1..9..80,M29163.1..9..76,
 M29164.1..9..117,9..173)
 /gene="PDHA1"
 /note="precursor"
 /codon_start=1
 /product="pyruvate dehydrogenase El-alpha subunit"
 /protein_id="AAA60051.1"
 /db_xref="GI:387009"
 /db_xref="GDB:G00-118-895"
 /translation="MRKMLAASRVLSGASQKPSASRVLSASRNPFANDATFEIKKCDLH
 RLEEGPPVTTLTREDGLKYRMQTVRMELKADQYKQKIIIRGFCHLDCGQACCV
 GLEAGINPTDHLITAYRAHGFTFTRGLSVREILAEITGRGGCAKGGKSHMYAKNF
 YGGNGIVGAQVPLGAGIALACKYNGKDEVCLTYGDGAANOQOIFAYNMAALWKLPC
 IFICENNRIGMGTVERAASTDYKKRDEFTPLGRVDMGMDILCVREATRFAAYCRSG
 KGPILMELOTYRVYHSHMSDPGVSYRTREIEQVRSKSDPFIMLLKDRVNSNLASVEE
 LKEIDVEVRKEIDAAQFATADPEPPELLEGIHYIYSSDPPEFVGANQWIKFKSVS"
 join(M29156.1..1..39..68,M29157.1..9..182,M29158.1..9..135,
 M29159.1..9..100,M29160.1..9..101,M29161.1..9..164,
 M29162.1..9..80,M29163.1..9..76,M29164.1..9..117,9..170)
 /gene="PDHA1"
 /note="G00-118-895"
 /product="pyruvate dehydrogenase El-alpha subunit"
 <1..8
 /gene="PDHA1"
 /note="G00-118-895"
 /number=10
 exon
 9..367
 /gene="PDHA1"
 /note="G00-118-895"
 /number=11
 polyA_signal
 342..347
 /gene="PDHA1"
 /note="G00-118-895"
 polyA_site
 359
 /gene="PDHA1"
 /note="G00-118-895"
 BASE COUNT 107 a 62 c 99 g 99 t
 ORIGIN About bp after segment 11; chromosome Xp22.1.

Query Match 100.0%; Score 273; DB 9; Length 367;
 Best Local Similarity 100.0%; Pred. No. 2.1e-66;
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ctaccacatctactccagcagccacccttttgaagttcgttggtgccaatcagtgatcaaa 60
 DB 95 CTACCACATCTACTCCAGCAGCCACCCTTTTGAAGTTCTGCTGGTCCCAATCAGTGATCAA 154
 QY 61 gtttaagtcagtcagttgaaggaggagagagaggttatcaccttcaggggggtaccag 120
 DB 155 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGAGGTTATACCTTCAGGGGGGTACCAG 214
 QY 121 acagtggttctcaacttgggttaaggaggagaaaacccagtcagtaattcaatgaaatt 180

Db 215 ACAGTGTTCTCAACTTGGTTAAGGAGGAGAAACCCAGTCAATGAAATTCGAATTAAT 274
 QY 181 ctggaacctccattgaagtgttagattgagcagcagtagtaattgcagtcagttgtaca 240
 Db 275 CTTGGAACATCTCCATTAAAGTGTGTAGATTGAGCAGGTAGTAGTAATTCATGTCACATGTTGTACA 334
 QY 241 ttagtcattaaagatgaattatttgagtgctt 273
 Db 335 TTAGTCATTAAAGATGAATATTATTGAGTGCTT 367
 RESULT 2
 LOCUS HUMPYDH 1410 bp mRNA linear PRI 27-APR-1993
 DEFINITION Human pyruvate dehydrogenase alpha subunit mRNA, complete cds.
 ACCESSION M24848
 VERSION M24848.1 GI:190761
 KEYWORDS pyruvate dehydrogenase alpha subunit.
 SOURCE Human liver, cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1410)
 AUTHORS Ho,L., Wexler,I.D., Liu,T.-C., Thekkumkara,T.J. and Patel,M.S.
 TITLE Characterization of cDNAs encoding human pyruvate dehydrogenase alpha-subunit
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86, 5330-5334 (1989)
 MEDLINE 89315791
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted by M.S.Patel, 16-MAY-1989
 FEATURES
 source
 Location/Qualifiers
 1. .1410
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 44..130
 sig_peptide
 /note="pyruvate dehydrogenase alpha subunit signal peptide"
 44..1216
 CDS
 /note="pyruvate dehydrogenase alpha subunit precursor (EC 1.2.4.1)"
 /codon_start=1
 /protein_id="AAA36533.1"
 /db_xref="GI:190762"
 /translation="MRKMLAASRVLSGASQKPSASRVLSASRNPFANDATFEIKKCDLH
 RLEEGPPVTTLTREDGLKYRMQTVRMELKADQYKQKIIIRGFCHLDCGQACCV
 GLEAGINPTDHLITAYRAHGFTFTRGLSVREILAEITGRGGCAKGGKSHMYAKNF
 YGGNGIVGAQVPLGAGIALACKYNGKDEVCLTYGDGAANOQOIFAYNMAALWKLPC
 IFICENNRIGMGTVERAASTDYKKRDEFTPLGRVDMGMDILCVREATRFAAYCRSG
 KGPILMELOTYRVYHSHMSDPGVSYRTREIEQVRSKSDPFIMLLKDRVNSNLASVEE
 LKEIDVEVRKEIDAAQFATADPEPPELLEGIHYIYSSDPPEFVGANQWIKFKSVS"
 131..1213
 mat_peptide
 /product="pyruvate dehydrogenase alpha subunit"
 BASE COUNT 373 a 296 c 402 g 339 t
 ORIGIN Unreported.
 Query Match 100.0%; Score 273; DB 9; Length 1410;
 Best Local Similarity 100.0%; Pred. No. 2e-66;
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ctaccacatctactccagcagccacccttttgaagttcgttggtgccaatcagtgatcaaa 60
 DB 1138 CTACCACATCTACTCCAGCAGCCACCCTTTTGAAGTTCTGCTGGTCCCAATCAGTGATCAA 1197
 QY 61 gtttaagtcagtcagttgaaggaggagagagaggttatcaccttcaggggggtaccag 120
 DB 1198 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGAGGTTATACCTTCAGGGGGGTACCAG 1257
 QY 121 acagtggttctcaacttgggttaaggaggagaaaacccagtcagtaattcaatgaaatt 180
 DB 1258 ACAGTGTTCTCAACTTGGTTAAGGAGGAGAAACCCAGTCAATGAAATTCGAATTAAT 1317

```

QY 181 cttggaattccattaaagtgttagattgagcagtaattgacgaattgttataca 240
Db 1318 CTTGGAACTTCATTAAAGTGTGTAGATTGACGAGTAGTAATTGTCAGTTGTGACA 1377
QY 241 ttatgtcattaaagatgaattattgagtgctt 273
Db 1378 TTAGTGCATTAAAGATGAATTATTGAGTGCTT 1410

RESULT 3
LOCUS HUMPDHA 1419 bp mRNA linear PRI 07-JAN-1995
DEFINITION Human pyruvate dehydrogenase alpha subunit mRNA, complete cds.
ACCESSION J03575
VERSION J03575.1 GI:189737
KEYWORDS pyruvate dehydrogenase.
SOURCE Human foreskin fibroblast, cDNA to mRNA, clone PHPDA.
ORGANISM Homo sapiens
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1419)
Koike,K., Ohta,S., Urata,Y., Kagawa,Y. and Koike,M.
Cloning and sequencing of cDNAs encoding alpha and beta subunits of
human pyruvate dehydrogenase
Proc. Natl. Acad. Sci. U.S.A. 85 (1), 41-45 (1988)
88124815
2 (bases 1 to 1419)
Koike,K., Urata,Y., Matsuo,S. and Koike,M.
Characterization and nucleotide sequence of human pyruvate
dehydrogenase alpha subunit gene
Unpublished (1990)
[2] revises [1].
Draft entry and printed copy of sequence for [1] kindly provided by
K.Koike, 11-APR-1988; and computer-readable copy for [2],
12-MAR-1990.

FEATURES
source
    Location/Qualifiers
    1..1419
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_type="fibroblast"
        /tissue_type="foreskin"
    <1..1419
        /gene="PDHA"
    1..1419
        /gene="PDHA"
    46..1218
        /gene="PDHA"
        /EC_number="1.2.4.1"
        /product="pyruvate dehydrogenase alpha subunit"
        /protein_id="AA060050.1"
        /db_xref="GI:189738"
        /translation="MRKMLAASRVLSGASQKSPASRVLSASRNFANDATFEIKKCDLH
        RLEEGPVTVTREDGLKYRRMOTVRRELKADOLYKOKIRGFCHLGDQOACCV
        GLEAGINPTDLHTIATYRAHGFYTRGLSVREILAEYTRGKGCAGKGGSMINAKNF
        YGFCENNRGMGTSPVRAAALCKYNGKDEVLTLYGDAANQGOIFEAYNNKALWKLPC
        IFTGNNRGMGTSPVRAAALCKYNGKDEVLTLYGDAANQGOIFEAYNNKALWKLPC
        KGPILMELQTYRHHGMSDPGVSYTRREIQVRSKSDPILMLKDRMYSNLSVVEE
        LKIDVVRKEIEDAAQAPADPEPPLEELGYHITSDDPPFVRGANQWIKFSVS"
    46..132
        /gene="PDHA"
    133..1215
        /gene="PDHA"
        /product="pyruvate dehydrogenase alpha subunit"
    377 a 237 c 403 g 342 t

BASE COUNT 377 a 237 c 403 g 342 t
ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 1419;
Best Local Similarity 100.0%; Pred. No. 2e-66; Indels 0; Gaps 0;
Matches 273; Conservative 0; Mismatches 0;

QY 1 ctaccacatctactccgcagccaccaccttttgagtcgtggtgagcgaatcagtgatcaa 60

```

```

Db 1140 CTACCACATCTACTCCAGCGACCCACCTTTTGAAGTTCGTGGTGCCAAATCAGTGATCAA 1199
QY 61 gtttaagtcaagtaaggaggaggaaggaggttatacccttcagggggctaccag 120
Db 1200 GTTAAAGTCAGTCAGTTAAAGGGGAGGAGAGAGGTTTACCTTCAGGGGGTACCAG 1259
QY 121 acagtgcttctcaacttggttaaggaggaagaaacccagcgaatgaatcaatgaatt 180
Db 1260 ACAGTGTCTCAACTTGGTTAAGGAGGAGAAACCCAGTCATGAATTCATGAATTT 1319
QY 181 cttggaattccattaaagtgttagattgagcaggtagtaattgacgaattgttataca 240
Db 1320 CTTGGAACTTCATTAAAGTGTGTAGATTGACGAGTAGTAATTGTCAGTTGTGACA 1379
QY 241 ttatgtcattaaagatgaattattgagtgctt 273
Db 1380 TTAGTGCATTAAAGATGAATTATTGAGTGCTT 1412

RESULT 4
LOCUS HUMPDHELB 1446 bp mRNA linear PRI 07-JAN-1995
DEFINITION Human pyruvate dehydrogenase El-alpha subunit mRNA, cds.
ACCESSION J03503
VERSION J03503.1 GI:189765
KEYWORDS pyruvate dehydrogenase.
SOURCE Human hepatoma, cDNA to mRNA, clone lambda-gt11.
ORGANISM Homo sapiens
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1446)
De Weirleir,L., Mackay,N., Lam Hon Wah,A.M. and Robinson,B.H.
Isolation of a full-length complementary DNA coding for human El
alpha subunit of the pyruvate dehydrogenase complex
J. Biol. Chem. 263 (4), 1991-1995 (1988)
88115327
COMMENT Clean copy of sequence [1] kindly provided by B.Robinson
(03/08/88).

FEATURES
source
    Location/Qualifiers
    1..1446
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
    <1..72
        /gene="PDHA1"
        /note="pyruvate dehydrogenase El-alpha signal peptide"
    <1..1245
        /gene="PDHA1"
        /note="pyruvate dehydrogenase El-alpha precursor"
        /codon_start=1
        /protein_id="AAA60055.1"
        /db_xref="GI:387011"
        /db_xref="GDB:G00-118-895"
        /translation="ETWGHPRRASVWVRSRRCHLCFMRKMLAASRVLSGASQKPA
        SRLVASRNFANDATFEIKKDLHREEGPVTVTREDGLKYRRMOTVRRELKADOLYKOKIRGFCHLGDQOACCV
        GLEAGINPTDLHTIATYRAHGFYTRGLSVREILAEYTRGKGCAGKGGSMINAKNF
        YGFCENNRGMGTSPVRAAALCKYNGKDEVLTLYGDAANQGOIFEAYNNKALWKLPC
        IFTGNNRGMGTSPVRAAALCKYNGKDEVLTLYGDAANQGOIFEAYNNKALWKLPC
        KGPILMELQTYRHHGMSDPGVSYTRREIQVRSKSDPILMLKDRMYSNLSVVEE
        LKIDVVRKEIEDAAQAPADPEPPLEELGYHITSDDPPFVRGANQWIKFSVS"
    46..132
        /gene="PDHA"
    133..1215
        /gene="PDHA"
        /product="pyruvate dehydrogenase El-alpha"
    379 a 307 c 414 g 346 t

BASE COUNT 379 a 307 c 414 g 346 t
ORIGIN 128 bp upstream of NaeI site.

Query Match 100.0%; Score 273; DB 9; Length 1446;
Best Local Similarity 100.0%; Pred. No. 2e-66; Indels 0; Gaps 0;
Matches 273; Conservative 0; Mismatches 0;

```

```
Qy 1 ctaccacatctactccagcgccaccaccccttttgaagttcgtggtgcccacatcagtggaatcaa 60
|||||
Db 1167 CTACCACATCTACTCCAGCGCCACCACCTTTTGAAGTTCGTGGTGCCCAATCAGTGGATCAA 1226
|||||
Qy 61 gtttaagtcagtcagtttaaggggaggaagagaggttatacccttcagggggctaccag 120
|||||
Db 1227 GTTTAAGTCAGTCAGTTAAGGGGAGGAGGAGGAGGTTATACCTTCAGGGGGCTACCAG 1286
|||||
Qy 121 acagtggtctcaactggtttaaggaggaagaaacccagtcgaatgaaattcaatgaaatt 180
|||||
Db 1287 ACAGTGTCTCAACTGTGTTAAGGAGGAAGAAACCAGTCAATGAATTCATGAATTT 1346
|||||
Qy 181 ctggaactccattaaagtgtagattgaagcaggttagtaattgcagttgttaca 240
|||||
Db 1347 CTTGGAACCTTCCATTAAAGTGTAGATTGACGAGGTAGTAATTGCATGTCAGTTGTACA 1406
|||||
Qy 241 ttatgcatataaagatgaattattgagtcgt 273
|||||
Db 1407 TTAGTGCATTAAGATGAATATTAGTGCTT 1439
|||||
..SULT 5
AX151747
LOCUS AX151747 1472 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 2 from Patent EP1118680.
ACCESSION AX151747
VERSION AX151747.1 GI:14533675
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1472)
AUTHORS Smith,J.C., Anand,R.A. and Morten,J.E.
TITLE Method for the diagnosis of polymorphisms in the human pdh elalpha
gene
JOURNAL Patent: EP 1118680-A 25-JUL-2001;
Astrazeneca AB (SE)
FEATURES
source 1. .1472
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 381 a 314 c 427 g 350 t
ORIGIN
Query Match 100.0%; Score 273; DB 6; Length 1472;
Best Local Similarity 100.0%; Pred. No. 2e-66;
atches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ctaccacatctactccagcgccaccaccccttttgaagttcgtggtgcccacatcagtggaatcaa 60
|||||
Db 1200 CTACCACATCTACTCCAGCGCCACCACCTTTTGAAGTTCGTGGTGCCCAATCAGTGGATCAA 1259
|||||
Qy 61 gtttaagtcagtcagtttaaggggaggaagagaggttatacccttcagggggctaccag 120
|||||
Db 1260 GTTTAAGTCAGTCAGTTAAGGGGAGGAGGAGGAGGTTATACCTTCAGGGGGCTACCAG 1319
|||||
Qy 121 acagtggtctcaactggtttaaggaggaagaaacccagtcgaatgaaattcaatgaaatt 180
|||||
Db 1320 ACAGTGTCTCAACTGTGTTAAGGAGGAAGAAACCAGTCAATGAATTCATGAATTT 1379
|||||
Qy 181 ctggaactccattaaagtgtagattgaagcaggttagtaattgcagttgttaca 240
|||||
Db 1380 CTTGGAACCTTCCATTAAAGTGTAGATTGACGAGGTAGTAATTGCATGTCAGTTGTACA 1439
|||||
Qy 241 ttatgcatataaagatgaattattgagtcgt 273
|||||
Db 1440 TTAGTGCATTAAGATGAATATTAGTGCTT 1472
|||||
```

RESULT 6

```
HUMPUDE1A
LOCUS 1472 bp mRNA linear PRI 08-JAN-1995
DEFINITION Human pyruvate dehydrogenase E1-alpha subunit (PDH1) mRNA,
complete cds.
ACCESSION L13318 J02734
VERSION L13318.1 GI:292412
KEYWORDS pyruvate dehydrogenase; pyruvate dehydrogenase E1-alpha subunit.
SOURCE Homo sapiens fetus and adult liver cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1472)
AUTHORS Dahl,H.H., Hunt,S.M., Hutchison,W.M. and Brown,G.K.
TITLE The human pyruvate dehydrogenase complex. Isolation of cDNA clones
for the E1 alpha subunit, sequence analysis, and characterization
of the mRNA
JOURNAL J. Biol. Chem. 262 (15), 7398-7403 (1987)
MEDLINE 87222349
REFERENCE 2 (sites)
AUTHORS Maragos,C., Hutchison,W.M., Hayasaka,K., Brown,G.K. and Dahl,H.H.
TITLE Structural organization of the gene for the E1 alpha subunit of the
human pyruvate dehydrogenase complex
JOURNAL J. Biol. Chem. 264 (21), 12294-12298 (1989)
MEDLINE 89308653
FEATURES
source 1. .1472
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
/dev_stage="fetus and adult"
1. .1472
/gene="PDH1"
1. .105
/gene="PDH1"
/feature="G00-118-895"
105..1278
/gene="PDH1"
/standard_name="pyruvate dehydrogenase complex"
/EC_number="1.2.4.1"
/codon_start=1
/product="pyruvate dehydrogenase E1-alpha subunit"
/protein_id="AAA60227.1"
/db_xref="GI:292413"
/db_xref="GDB:G00-118-895"
/translation="MRKMLAAVSRVLSGASQKSPASRVIVASRNFPANDATFIRKCDLH
RLEEGPVTTLTREDGLKYRMQTVRRMELADQLYKOKIIRGFCHLDCQDQACCV
GLEPAGINPTDHLITAYRAHGFTRGLSVREILAEITGRGGCAKGGKSHMYAKNF
YGGNGIVGQVPLGAGTALACKYNGDEVCITLYGCGAANQGFAYNRLAKLPC
IFICENNRQMGTSVERAAASTDYIKRGDITPLGRVDMIDILCVREATRYAAATCRSG
KGPLMELQTYRTHGSHSDPGSVSYRTREEIQEVRKSDPIMLLKDRWNSNLSVVEE
LKBIDIVEVRKEIEDAAQAFATADPEPPELLEELGYHYSDDPPFEVRGANQWIKFKGSV"
106..192
/gene="PDH1"
/feature="G00-118-895"
193..1275
/gene="PDH1"
/standard_name="pyruvate dehydrogenase complex"
/EC_number="1.2.4.1"
/feature="G00-118-895"
/product="pyruvate dehydrogenase E1-alpha subunit"
1276..1472
/gene="PDH1"
/feature="G00-118-895"
1447..1452
/gene="PDH1"
/feature="G00-118-895"
1472
/gene="PDH1"
/feature="G00-118-895"
381 a 314 c 427 g 350 t
BASE COUNT
ORIGIN
```


Mon Aug 19 09:13:23 2002

/db_xref="GI:219982"
/translation="MRKMLAAVSRVLSGASQKSPASRVLSVSRNFANDATFETKCDLH
RLEGPVTVLRTREGLKYRMQTVRRMELKADQLYKQKILRGCHLGDQACCV
GLEANGTDLHTAYRAHGFTTGLSVREILAEITGRGGCAKGGSMHVAKNE
YGGNGVAGPLGALACKYNGKDEVCLTLYGGAANOQOIFATYNNAAALWKLPC
IFICENNRNIGMTSVSRVAAASTDYKRGDFIPGLRVDMQDILCVREATRFPAAYCRSG
KGPILMELOQTYRTHGHSMSPGVSRTREIEQVRSKSDPIMLIKDRMVNSLASVEE
LKEIDVEVRKIEDAQAFAFADNPFPPEELGYHYIYSSDDPPFEVRGANQWIKFKSVS"
620..5840
/note="alpha subunit intron 1"
1818..2129
/note="Alu repeat 1"
5192..5479
/note="Alu repeat 2"
5841..5900
/note="alpha subunit exon 2"
5901..6371
/note="alpha subunit intron 2"
6372..6545
/note="alpha subunit exon 3"
6546..7708
/note="alpha subunit intron 3"
6677..6999
/note="Alu repeat 3"
7709..7835
/note="alpha subunit exon 4"
7836..9499
/note="alpha subunit intron 4"
8223..8517
/note="Alu repeat 4"
8642..8956
/note="Alu repeat 5"
9500..9591
/note="alpha subunit exon 5"
9592..10907
/note="alpha subunit intron 5"
10387..10648
/note="Alu repeat 6"
10908..11000
/note="alpha subunit exon 6"
11001..11766
/note="alpha subunit intron 6"
11767..11922
/note="alpha subunit exon 7"
11923..12106
/note="alpha subunit intron 7"
12107..12178
/note="alpha subunit exon 8"
12179..14063
/note="alpha subunit intron 8"
13559..13809
/note="Alu repeat 7"
14064..14131
/note="alpha subunit exon 9"
14132..15317
/note="alpha subunit intron 9"
15318..15426
/note="alpha subunit exon 10"
15427..15887
/note="alpha subunit intron 10"
15888..16253
/note="alpha subunit exon 11"
16221..16226
polyA_signal
BASE COUNT 4377 a 3680 c 3957 g 5058 t
ORIGIN

Query Match 100.0%; Score 273; DB 9; Length 17072;
Best Local Similarity 100.0%; Pred. NO. 1.8e-66;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ctaccacatctactccagcagccacccttttgagttcgtggtgccaatcagtgatcaaa 60
|||||

Query Match 100.0%; Score 273; DB 9; Length 1472;
Best Local Similarity 100.0%; Pred. No. 2e-66;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ctaccacatctactccagcagccacccttttgagttcgtggtgccaatcagtgatcaaa 60
Db - 1200 CTACCACATCTACTCCAGCAGCCACCCTTTTGAAGTTCTGTTGCGCAATCAGTGGATCAA 1259
Oy 61 gtttaagtcagtcagtttaagggaggagaggaggttatcccttcaggggctaccag 120
Db 1260 GTTTAAGTCAGTCAGTTAAGGGAGGAGAGAGAGGTTATACCTTCAGGGGCTACCA 1319
Oy 121 acagtggttcaactgttgaagagagaagaaacccagtcgaatgaattcaatgaatt 180
Db 1320 ACAGTGTTCACACTTGTGTGAAGAGGAGAGAGAAACCCAGTCGAATGAATTCATGAAT 1379
Oy 181 ctggaaacttcattagtgtagattgagcaggtagtaattgacatgagttgtaca 240
Db 1380 CTTGGAACCTCCATTAGTGTGTAGATTGACAGGAGTAGTAATGCGATGCGATTTGTACA 1439
241 tttagtcattaaagatgaattatgagtgctt 273
Db 1440 TTAGTGCAATAAAGATGATTTATTGAGTGCTT 1472
RESULT 7
HUMPDHAL 17072 bp DNA linear PRI 01-FEB-2000
LOCUS Human pyruvate dehydrogenase [EC 1.2.4.1] alpha subunit gene, exons
DEFINITION 1-11.
ACCESSION D90084 M58568
VERSION D90084.1 GI:219981
KEYWORDS pyruvate dehydrogenase; pyruvate dehydrogenase alpha subunit.
SOURCE Human leukocyte DNA, clones pGPDHA[13 and 37].
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 17072)
Koike,K., Urata,Y., Matsuo,S. and Koike,M.
Characterization and nucleotide sequence of the gene encoding the
human pyruvate dehydrogenase alpha-subunit
Gene 93 (2), 307-311 (1990)
JOURNAL 91033044
MEDLINE These data kindly submitted in computer readable form by: Kichiko
COMMENT Koike
Department of Pathological Biochemistry
Atomic Disease Institute Nagasaki University School of Medicine
12-4, Sakamoto-machi
Nagasaki-shi 852
Japan
Phone: 0958-47-2111
Fax: 0958-47-8514
First 10 bp reported in [1] is linker sequence.
FEATURES
source 1..17072
/organism="Homo sapiens"
/db_xref="taxon:9606"
protein_bind 286..291
/bound_moiety="SpI"
CAAT_signal 342..346
protein_bind 370..375
/bound_moiety="SpI"
TATA_signal 429..434
prim_transcript 439..16253
/note="alpha subunit mRNA and introns"
exon 439..619
/note="alpha subunit exon 1"
CDS join(563..619,5841..5900,6372..6545,7709..7835,9500..10908,
11000,11767..11922,12107..12178,14064..14131,
15318..15426,15888..16052)
/note="alpha subunit"
/codon_start=1
/protein_id="BAAL14121.1"

us-09-761-579-2_copy_1200_1472.rge

Mon Aug 19 09:13:23 2002

```

/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B-8"
/notes="Vector: pOTB7"
58.1230
/codon_start=1
/product="pyruvate dehydrogenase (lipoamide) alpha 1"
/protein_id="AAH02406.1"
/db_xref="GI:12803199"
/translatlon="MRKMLAAVSRVLSGASQSPASRVLVASRNFANDATFEKKDLH
RLEGEPVTVLTREDGLKYRMQTVRMELKADQLYKQIIRGFCHECDGQACCV
GLEAGNPTDLHITAYRAGFTFTRGLSVREILAEITLTKRGCGKAGKSGMHYAKNF
YGGNIVGAQVPLGATLACKYNGDEVLTLVGDGAANOQOIFAYNMAALWKLPC
IFICENNRGMGTSGVERAAADSYTKRGEIPGLRVGDMILCVREATRFPAAYCRSG
KGPLMELOTYRHYHSMSPDYSYTRREIQRVRSKSDPIMLLKDMVNSNLASVEE
LKEIDVEVRKEIEDAAQFATADPEPPELGYHIYSDDPPFEVRGANQWIKFKSVS"
BASE COUNT 408 a 302 c 406 g 340 t
ORIGIN
Query Match 98.9%; Score 270; DB 9; Length 1456;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ctacacatctactccagcgacccaccttttgaagtctggtggtgccaatcagtgatcaa 60
DB 1152 CTACCACTCTACTCCAGCGACCCACCTTTTGAAGTTCGTGTGCGCAATCAGTGGATCAA 1211
QY 61 gtttaagtcagtcagtaagggaggagagagagagagagagagagagagagagagag 120
DB 1212 GTTTAAGTCAGTCAGTTAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1271
QY 121 acagtgcttcaactggttaagggaggagagagagagagagagagagagagagagag 180
DB 1272 ACAGTGTCTCACTGTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1331
QY 181 cttggaacttcattgaagtgtagtagtagtagtagtagtagtagtagtagtagtagtag 240
DB 1332 CTTGGAACCTTCCATTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1391
QY 241 ttagtcattaaaagatgaattattgagt 270
DB 1392 TTAGTCATTAAAGATGAATTATTGAGTG 1421

RESULT 11
HUMPHALG HUMPHALG .1492 bp mRNA linear PRI 15-AUG-2000
LOCUS Homo sapiens pyruvate dehydrogenase E1-alpha subunit precursor
DEFINITION (PDHAL) mRNA, complete cds.
ACCESSION L48690
VERSION L48690.1 GI:1051096
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1492)
AUTHORS Dahl, H.H., Brown, G.K., Brown, R.M., Hansen, L.L., Kerr, D.S.,
Wexler, I.D., Patel, M.S., De Meirleir, L., Lissens, W., Chun, K.,
Mackay, N. and Robinson, B.H.
Mackay, N. and Robinson, B.H.
Mutations and polymorphisms in the pyruvate dehydrogenase E1 alpha
gene
Hum. Mutat. 1 (2), 97-102 (1992)
JOURNAL MEDLINE 93244853
PUBMED 1301207
REFERENCE 2 (bases 1 to 1492)
AUTHORS Takakubo, F., Cartwright, P., Hoogenraad, N., Thorburn, D.R.,
Collins, F., Lichtig, T., and Dahl, H.H.
An amino acid substitution in the pyruvate dehydrogenase E1 alpha
gene, affecting mitochondrial import of the precursor protein
Am. J. Hum. Genet. 57 (4), 772-780 (1995)
JOURNAL MEDLINE 96029268

```

```

Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ctacacatctactccagcgacccaccttttgaagtctggtggtgccaatcagtgatcaa 60
DB 1149 CTACCACTCTACTCCAGCGACCCACCTTTTGAAGTTCGTGTGCGCAATCAGTGGATCAA 1208
QY 61 gtttaagtcagtcagtaagggaggagagagagagagagagagagagagagagagag 120
DB 1209 GTTTAAGTCAGTCAGTTAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1268
QY 121 acagtgcttcaactggttaagggaggagagagagagagagagagagagagagagag 180
DB 1269 ACAGTGTCTCACTGTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1328
QY 181 cttggaacttcattgaagtgtagtagtagtagtagtagtagtagtagtagtagtagtag 240
DB 1329 CTTGGAACCTTCCATTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1388
QY 241 ttagtcattaaaagatgaattattgagt 270
DB 1389 TTAGTCATTAAAGATGAATTATTGAGTG 1418

RESULT 10
BC002406 1456 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, pyruvate dehydrogenase (lipoamide) alpha 1, clone
DEFINITION MGC:8609 IMAGE:2961286, mRNA, complete cds.
ACCESSION BC002406
VERSION BC002406.1 GI:12803198
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1456)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgpb@remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shcherbakov, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantropop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 1 Row: 1 Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505684.
Location/Qualifiers
1. 1456
/organism="Homo sapiens"
/db_xref="LocusID:5160"
/db_xref="taxon:9606"
/clone="MGC:8609 IMAGE:2961286"
FEATURES
source

```

pyruvate dehydrogenase.

KEYWORDS
SOURCE
ORGANISM
AUTHORS
REFERENCE
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

Human
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria, Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1417)
Song, B.J.
Direct Submission
Submitted (18-APR-1990) Song B.J., National Institute on Alcohol
Abuse and Alcoholism, Laboratory of Metabolism and Molecular
Biology, 12501 Washington Ave., Rockville, MD 20852, USA
2 (bases 1 to 1417)
Hub, T.L., Chi, Y.T., Casazza, J.P., Veech, R.L. and Song, B.J.
Unpublished
Unpublished (05-NOV-1990) by Song B.J.

```

COMMENT
FEATURES
Data kind/Qualifiers
1. .1417
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="Xpter-q13."
/clone.lib="liver in lambda gtl1."
57. .1229
/note="pyruvate dehydrogenase E1' subunit (AA 1 - 390)"
/codon_start=1
/protein_id="CAA36934.1"
/db_xref="GI:35381"
/db_xref="SWISS-PROT:P08559"
/translation="MRKMLAAVSRLSCASQPSRVLVASRNFANDATFEIKKCDLH
RIEUGPVTVTLTRDGLKUYRMTQVTVRMELKADQLYKQKIRGFCGLCQDQACCV
GLEAGINTDHLITRAYRHGHTFTKGLSVREILAEITGKGGCKGKSGSHMYAKLP
YGCNIGVAQPLGVAGIALCKYKNGDDEVCLTILGDGAANQOIFEANYMALAKLPC
ITFCENRRTGMTSVRAAASTDYYIKRGDFIPGLRVGSDMDILCVKRTAFNAYCRSG
KGPILMEQLTQYTHGHSMDSPGVSYRTREETQEVRSKSDPIMLLKDRMTVNSLWAVEE
KLKVIDEVKKEITDAQATQADPEPPLGELGYHYISDPPFPFVRGANOWIKPKFSV"
373 a 302 c 404 g 338 t

```

BASE COUNT	373 a	302 c	404 g	338 t
ORIGIN				
Query Match		97.8%	Score 267;	DB 9; Length 1417;
Best Local Similarity		100.0%	Pred. No. 9, 6e-65;	
Insertions		0;	Mismatches	0; Indels 0; Gaps 0;

QY	1	ctacacacatctactccagcaccacacctttgaagtcttggtgcccgaattcagtggatcaa	60
DB	1151	CTACACATCTACTCCAGCGACCACCTTTTGAAGTCGTGGTGCCANTCAGTGGATCAA	1210
QY	61	gtttaagtcagtcagtttaagggggagaagagattataocttcagggggctaccag	120
DB	1211	GTTTTAAGTCAGTCAGTTAAAGGGGAGGAGGAGGGTTTATCCTTCAGGGGGCTACCAG	1270
QY	121	acagtggttctcaactgggttaaaggaggaagaaaccacagtcgaatgaattcaaatgaatt	180
DB	1271	ACAGTGTGTTCTCAACTTGGTTAAGGAGGAGAANAACCCAGTCAGTGAATTCANATGAANAATT	1330
QY	181	cttggaaaacttcocatgaagtgtagattgagcaggtagtaattgcagcttgtttgaca	240
DB	1331	CITGGAAAATCTCCATTAAGTGTGTAGATTGACAGGTAGTAGTTGCATGCCAGTTGTGTACA	1390
QY	241	ttagtgcatataaagatgaattatga	267
DB	1391	TTAGTGCATATAAAGATGAATTAATTGA	1417

RESULT	13
GLI3505	
LOCUS	GLI3505 STS SHGC-10981, sequence tagged site.
DEFINITION	human STS SHGC-10981, sequence tagged site.
ACCESSION	GI3505.1 GI:1129244
VERSION	STG; STS sequence; primer; sequence tagged site.
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Chordata; Craniata; Vertebrata; Euteleostomi;

```

PUBMED 7573035
FEATURES
    source
        Location/Qualifiers
            1..1492
                /organism="Homo sapiens"
                /isolate="920002"
                /db_xref="taxon:9606"
                /sex="male"
                /cell_type="fibroblast"
                /tissue_type="skin"
                /dev_stage="ten month old"
            <1..1492
                /gene="PDHAL"
                106..1278
                /gene="PDHAL"
                /codon_start=1

```

product="pyruvate dehydrogenase E1-alpha subunit"	precursor	protein_id="AAB59581.1"	db_xref="GI:1051097"	translation="MRKMLAANSPVLGGASQRPASRLVLSNFANDATFEIKKCDLH RLEGPPVTLVLRDEGLQYKRMQTVRRMELKADQLYKOKIIRGCHICDQAEACV GLRASPQTIISSQPTGLTALLSPGAFPEKFSQSLQDEKEVVLKGGEDCTCMPTPS TGAMASNERCPWAAGIALACKYNGKDEVCLTYLGDGAANQQOIFEAYNMAALMLPC YICENNRYGMGTSGVERAAATDYIKRGDFIPLGRVDGMDILCVREATFAAAYCRSG KGPILMELQTVRHGSHSSDPGVSYRTREIQEVRSKSDPIMLLKDRMYSNLSLAQEE LKEIDVFRKEIEDAAQAFATADPEPPELLELGYHIYSSDPPFEVRGANQWIKPKFSV"
	sig_peptide	106..192	/gene="PDHAL"	/note="G00-118-895"
	variation	134	/note="G00-118-895"	
		134	/gene="PDHAL"	/note="results in arginine to proline substitution"
	mat_peptide	193..1275	/replace="g"	
		193..1275	/gene="PDHAL"	
		193..1275	/note="G00-118-895"	
		193..1275	/product="pyruvate dehydrogenase E1-alpha subunit"	

Query Match	98.2%	Score 268.2	DB 9	Length 1492	
Best Local Similarity	98.9%	Pred. No. 4.4e-65			
Matches 270	Conservative	0	Mismatches 3	Indels 0	Gaps 0
QY	1	ctacacatctactcagcgaccacccctttgaagttcgtgtgccaatcagtgatcaa	60		
Y	1200	CTACACATCTACTCCAGGCCACCCCTTTGAAGTTCGTGTGGTGGCAATCAGTGGATCAA	1259		
AY	61	gtttaagtcagtcagtttaaggggagagagagaggttataccttcaggggggtaccag	120		
DB	1260	GTTTAAAGTCAGTCAGTTAAAGGGGAGGAGAGAGGTTATACCTTCACGGGGGCTACCAAG	1319		
QY	121	acagtggttcacacttggttaagagagaagaacccagccaatgaattcaatgaaatt	180		
DB	1320	ACRGTTGTTCTCACTTGGTTAAGGAGGAGCAAAAACCCAGTCGAATGAATTCGAATGAAT	1379		
QY	181	cttggaaaacttcocattaaagtggttagattagcagcaggtagtaattgatcagttgtaca	240		
DB	1380	CTTGAAAACCTCCCATTAAGTGTGTAGATTGAGCAGGATGACATGTCAGTGGTTGTACA	1439		
QY	241	ttagtgcatataaagatgaattatttgagtgtt	273		
Y	1440	TTAGTGCAATTAAGATCAATCAATTAATGAGTGCTT	1472		

RESULT	12
HSPDHLE1	
LOCUS	HSPDHLE1
DEFINITION	Human mRNA for liver pyruvate dehydrogenase (EC 1.2.4.1) E1,
	subunit.
	1417 bp
	mRNA
	linear
	PRI 12-SEP-1993
ACCESSION	X52710
VERSION	X52710.1
	GI:35380

Mon Aug 19 09:13:23 2002

TITLE Molecular genetic analysis of a female patient with pyruvate dehydrogenase deficiency: detection of a new mutation and differential expression of mutant gene product in cultured cells J. Inherit. Metab. Dis. 18 (5), 547-557 (1995)

JOURNAL 96170879

MEDLINE

REMARK GenBank staff at the National Library of Medicine created this entry (NCBI gisbq 176881) from the original journal article. This sequence comes from Fig. 4.

COMMENT Map location: Xp22.1-22.2.

FEATURES

source

1..149

/organism="Homo sapiens"

/organelle="mitochondrion"

/db_xref="taxon:9606"

gene

1..149

/partial

/gene="pyruvate dehydrogenase complex alpha subunit, PDHC, E1 alpha"

CDS

1..138

/partial

/gene="pyruvate dehydrogenase complex alpha subunit, PDHC, E1 alpha"

/note="This sequence comes from Fig. 4; Map location Xp22.1-22.2; PDHC; E1 alpha"

/codon_start=1

/product="pyruvate dehydrogenase complex alpha subunit"

/protein_id="AAB36217.1"

/db_xref="GI:1478430"

/translation="FATADPEPPEDEPPELGYHYSSDPPFVYRGANOWIKFKSV"

BASE COUNT 37 a 37 c 40 g 35 t

ORIGIN

Query Match 33.0%; Score 90; DB 9; Length 149;

Best Local Similarity 100.0%; Pred. No. 6.8e-15;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaccacatctactccagcagccaccttttgaagttcgtggtgccaatcagtgatcaa 60

Db 60 CTACCACATCTACTCCAGCGACCCACTTTTGAGTTCGTGTGCGCAATCAGTGATCAA 119

QY 61 gtttaagtcagtcagtttaaggagagagaa 90

Db 120 GTTTAAGTCAGTCAGTTAAGGGGAGGAGAA 149

Search completed: August 16, 2002, 13:40:31

Job time: 6830 sec